

Complex ALLO

Diego Gomez

January 2025

1 Stability analysis

1.1 Notation

Let $\mathbf{L} \in \mathbb{R}^{|S| \times |S|}$ be a finite dimensional matrix whose eigendecomposition we are interested in approximating. Let $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^\top$ be its eigendecomposition, with $\mathbf{U} = [\mathbf{u}_i]_{i=1}^d$ and $\mathbf{V} = [\mathbf{v}_i]_{i=1}^d$ the matrices of left and right eigenvectors, and $\mathbf{\Lambda} = \text{diag}([\lambda_i]_{i=1}^d)$ the diagonal matrix containing the eigenvalues. We will not make assumptions about \mathbf{L} being symmetric, which means that $\mathbf{u}_i, \mathbf{v}_i \in \mathbb{C}^{|S|}$ and $\lambda_i \in \mathbb{C}$, i.e., the eigensystem of \mathbf{L} is complex. For ease of notation, we denote any complex quantity x as $x = a^x + ib^x$.

Further, we will use $X = \{\mathbf{x}_i\}_{i=1}^d$ and $Y = \{\mathbf{y}_i\}_{i=1}^d$ to denote the sets of vectors approximating the left and right eigenvectors. Also, let us introduce the constraint errors $h_{x,jk} = \langle \bar{\mathbf{x}}_j, \llbracket \mathbf{y}_k \rrbracket \rangle_\rho - \delta_{jk}$ and $h_{y,jk} = \langle \llbracket \bar{\mathbf{x}}_k \rrbracket, \mathbf{y}_j \rangle_\rho - \delta_{jk}$, where $\rho \in \Delta(\mathcal{S})$ is a state distribution of choice. In particular, we can write the constraint errors as: $h_{x,jk} = \mathbf{x}_j^\top \mathbf{D} \llbracket \mathbf{y}_k \rrbracket - \delta_{jk}$ and $h_{y,jk} = \llbracket \mathbf{x}_k \rrbracket^\top \mathbf{D} \mathbf{y}_j - \delta_{jk}$, with $\mathbf{D} = \text{diag}(\rho)$.

When the approximations are exactly equal to the eigenvectors, the errors become 0, meaning that their real part is equal to the identity matrix \mathbf{I} and the imaginary one, to the null matrix \mathbf{Z} . We associate to the real errors the dual variables $\boldsymbol{\alpha}^x = [\alpha_{x,jk}]_{1 \leq k \leq j \leq d}^d$ and $\boldsymbol{\alpha}^y = [\alpha_{y,jk}]_{1 \leq k \leq j \leq d}^d$, and to the imaginary ones, $\boldsymbol{\beta}^x = [\beta_{x,jk}]_{1 \leq k \leq j \leq d}^d$ and $\boldsymbol{\beta}^y = [\beta_{y,jk}]_{1 \leq k \leq j \leq d}^d$.

The squared augmented Lagrangian objective takes the form:

$$\begin{aligned} \mathcal{L}(X, Y, \boldsymbol{\alpha}^x, \boldsymbol{\alpha}^y, \boldsymbol{\beta}^x, \boldsymbol{\beta}^y) = & \sum_{i=1}^d \frac{1}{2} (\text{Re}(\langle \bar{\mathbf{x}}_i, \mathbf{L} \mathbf{y}_i \rangle_\rho))^2 - \sum_{j=1}^d \sum_{k=1}^j \alpha_{x,jk} \text{Re}(h_{x,jk}) + \frac{b}{2} \sum_{j=1}^d \sum_{k=1}^j (\text{Re}(h_{x,jk}))^2 \cdots \\ & \cdots - \sum_{j=1}^d \sum_{k=1}^j \alpha_{y,jk} \text{Re}(h_{y,jk}) + \frac{b}{2} \sum_{j=1}^d \sum_{k=1}^j (\text{Re}(h_{y,jk}))^2 - \sum_{j=1}^d \sum_{k=1}^j \beta_{x,jk} \text{Im}(h_{x,jk}) \cdots \\ & \cdots + \frac{b}{2} \sum_{j=1}^d \sum_{k=1}^j (\text{Im}(h_{x,jk}))^2 - \sum_{j=1}^d \sum_{k=1}^j \beta_{y,jk} \text{Im}(h_{y,jk}) + \frac{b}{2} \sum_{j=1}^d \sum_{k=1}^j (\text{Im}(h_{y,jk}))^2. \end{aligned}$$

1.2 Dynamical system

Since we are using the dual framework, we have a dynamical system where each \mathbf{x}_i and \mathbf{y}_i is changing to minimize \mathcal{L} and, in opposition, each dual is changing to maximize \mathcal{L} . To make more explicit the behavior of the resulting system, let us calculate the gradients driving the dynamics. First, for the real component of \mathbf{x}_i :

$$\begin{aligned} \mathbf{g}_{\mathbf{a}_i^x} &:= \nabla_{\mathbf{a}_i^x} \mathcal{L} = \text{Re}(\langle \bar{\mathbf{x}}_i, \mathbf{L}\mathbf{y}_i \rangle_\rho) \nabla_{\mathbf{a}_i^x} \text{Re}(\langle \bar{\mathbf{x}}_i, \mathbf{L}\mathbf{y}_i \rangle_\rho) - \sum_{k=1}^i (\alpha_{x,ik} - b \cdot \text{Re}(h_{x,ik})) \text{Re}(\nabla_{\mathbf{a}_i^x} h_{x,ik}) \\ &\quad - \sum_{k=1}^i (\beta_{x,ik} - b \cdot \text{Im}(h_{x,ik})) \text{Im}(\nabla_{\mathbf{a}_i^x} h_{x,ik}); \end{aligned}$$

$$\begin{aligned} \nabla_{\mathbf{a}_i^x} \text{Re}(\langle \bar{\mathbf{x}}_i, \mathbf{L}\mathbf{y}_i \rangle_\rho) &= \nabla_{\mathbf{a}_i^x} \text{Re}((\mathbf{a}^{\mathbf{x}_i} + i\mathbf{b}^{\mathbf{x}_i})^\top \mathbf{D}\mathbf{L}(\mathbf{a}^{\mathbf{y}_i} + i\mathbf{b}^{\mathbf{y}_i})) \\ &= \nabla_{\mathbf{a}_i^x} \left[(\mathbf{a}^{\mathbf{x}_i})^\top \mathbf{D}\mathbf{L}\mathbf{a}^{\mathbf{y}_i} - (\mathbf{b}^{\mathbf{x}_i})^\top \mathbf{D}\mathbf{L}\mathbf{b}^{\mathbf{y}_i} \right] \\ &= \mathbf{D}\mathbf{L}\mathbf{a}^{\mathbf{y}_i}; \end{aligned}$$

$$\begin{aligned} \nabla_{\mathbf{a}_i^x} h_{x,jk} &= \nabla_{\mathbf{a}_i^x} \mathbf{x}_j^\top \mathbf{D}[\mathbf{y}_k] \\ &= \nabla_{\mathbf{a}_i^x} (\mathbf{a}^{\mathbf{x}_j} + i\mathbf{b}^{\mathbf{x}_j})^\top \mathbf{D}[(\mathbf{a}^{\mathbf{y}_k} + i\mathbf{b}^{\mathbf{y}_k})] \\ &= \delta_{ij} \mathbf{D}(\mathbf{a}^{\mathbf{y}_k} + i\mathbf{b}^{\mathbf{y}_k}); \end{aligned}$$

$$\begin{aligned} \therefore \mathbf{g}_{\mathbf{a}_i^x} &= \text{Re}(\langle \bar{\mathbf{x}}_i, \mathbf{L}\mathbf{y}_i \rangle_\rho) \mathbf{D}\mathbf{L}\mathbf{a}^{\mathbf{y}_i} - \sum_{k=1}^i (\alpha_{x,ik} - b \cdot \text{Re}(h_{x,ik})) \mathbf{D}\mathbf{a}^{\mathbf{y}_k} \\ &\quad - \sum_{k=1}^i (\beta_{x,ik} - b \cdot \text{Im}(h_{x,ik})) \mathbf{D}\mathbf{b}^{\mathbf{y}_k}. \end{aligned}$$

Similarly, we have for the imaginary component:

$$\begin{aligned} \mathbf{g}_{\mathbf{b}_i^x} &= -\text{Re}(\langle \bar{\mathbf{x}}_i, \mathbf{L}\mathbf{y}_i \rangle_\rho) \mathbf{D}\mathbf{L}\mathbf{b}^{\mathbf{y}_i} + \sum_{k=1}^i (\alpha_{x,ik} - b \cdot \text{Re}(h_{x,ik})) \mathbf{D}\mathbf{b}^{\mathbf{y}_k} \\ &\quad - \sum_{k=1}^i (\beta_{x,ik} - b \cdot \text{Im}(h_{x,ik})) \mathbf{D}\mathbf{a}^{\mathbf{y}_k} \end{aligned}$$

Together, we have that the steepest ascent direction with respect to \mathbf{x}_i is:

$$\mathbf{g}_{\mathbf{x}_i} = \mathbf{g}_{\mathbf{a}_i^x} + i\mathbf{g}_{\mathbf{b}_i^x}$$

$$\begin{aligned}
&= \text{Re}(\langle \bar{\mathbf{x}}_i, \mathbf{L}\mathbf{y}_i \rangle_\rho) \mathbf{D}\mathbf{L}(\mathbf{a}^{\mathbf{y}_i} - i\mathbf{b}^{\mathbf{y}_i}) - \sum_{k=1}^i (\alpha_{x,ik} - b \cdot \text{Re}(h_{x,ik})) \mathbf{D}(\mathbf{a}^{\mathbf{y}_k} - i\mathbf{b}^{\mathbf{y}_k}) \\
&\quad - \sum_{k=1}^i (\beta_{x,ik} - b \cdot \text{Im}(h_{x,ik})) \mathbf{D}(\mathbf{b}^{\mathbf{y}_k} + i\mathbf{a}^{\mathbf{y}_k}) \\
&= \text{Re}(\langle \bar{\mathbf{x}}_i, \mathbf{L}\mathbf{y}_i \rangle_\rho) \mathbf{D}\mathbf{L}\bar{\mathbf{y}}_i - \sum_{k=1}^i (\alpha_{x,ik} - b \cdot \text{Re}(h_{x,ik})) \mathbf{D}\bar{\mathbf{y}}_k \\
&\quad + i \sum_{k=1}^i (\beta_{x,ik} - b \cdot \text{Im}(h_{x,ik})) \mathbf{D}(i\mathbf{b}^{\mathbf{y}_k} - \mathbf{a}^{\mathbf{y}_k}) \\
&= \text{Re}(\langle \bar{\mathbf{x}}_i, \mathbf{L}\mathbf{y}_i \rangle_\rho) \mathbf{D}\mathbf{L}\bar{\mathbf{y}}_i - \sum_{k=1}^i (\alpha_{x,ik} + i\beta_{x,ik} - bh_{x,ik}) \mathbf{D}\bar{\mathbf{y}}_k.
\end{aligned}$$

If we repeat the same process for the right approximation \mathbf{y}_i , we will obtain a very similar expression, except that every matrix is now transposed and \mathbf{x}_k replaces \mathbf{y}_k :

$$\mathbf{g}_{\mathbf{y}_i} = \mathbf{g}_{\mathbf{a}_i^{\mathbf{y}}} + i\mathbf{g}_{\mathbf{b}_i^{\mathbf{y}}} = \text{Re}(\langle \bar{\mathbf{x}}_i, \mathbf{L}\mathbf{y}_i \rangle_\rho) \mathbf{L}^\top \mathbf{D}\bar{\mathbf{x}}_i - \sum_{k=1}^i (\alpha_{y,ik} + i\beta_{y,ik} - bh_{y,ik}) \mathbf{D}\bar{\mathbf{x}}_k.$$

Additionally, the gradient with respect to the duals is equal to their corresponding constraint error term:

$$\mathbf{g}_{\alpha_{x,jk}} = \text{Re}(h_{x,jk}) ; \quad \mathbf{g}_{\alpha_{y,jk}} = \text{Re}(h_{y,jk}) ; \quad \mathbf{g}_{\beta_{x,jk}} = \text{Im}(h_{x,jk}) ; \quad \mathbf{g}_{\beta_{y,jk}} = \text{Im}(h_{y,jk}).$$

For ease of notation, let us denote $\Theta[t] = \text{vec}\{X[t], Y[t], \boldsymbol{\alpha}^x[t], \boldsymbol{\alpha}^y[t], \boldsymbol{\beta}^x[t], \boldsymbol{\beta}^y[t]\}$ the vectorized form of the primal and dual parameters. Also, let

$$\mathbf{g}_\Theta = \text{vec}\left\{ \bigcup_i \mathbf{g}_{\mathbf{x}_i}, \bigcup_i \mathbf{g}_{\mathbf{y}_i}, \bigcup_{jk} -\mathbf{g}_{\alpha_{x,jk}}, \bigcup_{jk} -\mathbf{g}_{\alpha_{y,jk}}, \bigcup_{jk} -\mathbf{g}_{\beta_{x,jk}}, \bigcup_{jk} -\mathbf{g}_{\beta_{y,jk}} \right\}$$

be the corresponding vectorized gradients. Our dynamical system is given by:

$$\Theta[t+1] = \Theta[t] - \alpha \cdot \mathbf{g}_\Theta(\Theta[t]).$$

1.3 Equilibria

To find the equilibrium points of the system we just need to solve for the dynamics becoming stationary, i.e., when the gradients become 0.

Trivially, requiring $g_{\alpha_{x,jk}} = g_{\alpha_{y,jk}} = g_{\beta_{x,jk}} = g_{\beta_{y,jk}} = 0$ implies that the biorthogonality constraints are satisfied. For $\mathbf{g}_{\mathbf{x}_i}$ and $\mathbf{g}_{\mathbf{y}_i}$ to be $\mathbf{0}$, we must have:

$$\text{Re}(\langle \bar{\mathbf{x}}_i^*, \mathbf{L}\mathbf{y}_i^* \rangle_\rho) \mathbf{D}\mathbf{L}\bar{\mathbf{y}}_i^* = \sum_{k=1}^i (\alpha_{x,ik}^* + i\beta_{x,ik}^*) \mathbf{D}\bar{\mathbf{y}}_k^*$$

$$\text{Re}(\langle \bar{\mathbf{x}}_i^*, \mathbf{L} \mathbf{y}_i^* \rangle_\rho) \mathbf{L}^\top \mathbf{D} \bar{\mathbf{x}}_i^* = \sum_{k=1}^i (\alpha_{y,ik}^* + i \beta_{y,ik}^*) \mathbf{D} \bar{\mathbf{x}}_k^*.$$